AD_____

GRANT NUMBER: DAMD17-94-J-4270

TITLE: A New Generic Method for the Production of Protein-Based Inhibitors of Proteins Involved in Cancer Metastasis

PRINCIPAL INVESTIGATOR: Marshall H. Edgell, Ph.D.

CONTRACTING ORGANIZATION: The University of North Carolina at Chapel Hill
Chapel Hill, North Carolina 27599-4100

REPORT DATE: July 31, 1995

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 31, 1995	Annual 1 Aug 94 - 31 Jul 95

4. TITLE AND SUBTITLE
A New Generic Method for the Production of Protein-Based Inhibitors of Proteins Involved in Cancer Metastasis

5. FUNDING NUMBERS
DAMD17-94-J-4270

6. AUTHOR(S)
Marshall H. Edgell, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of North Carolina at Chapel Hill
Chapel Hill, North Carolina 27599-4100

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Metastasis is a complex, multi-step process involving numerous proteins. Our long-range interest is to design protein-based inhibitors that will block various steps in metastasis. As our first target we have chosen stromelysin, a proteinase implicated in metastasis. There is considerable evidence implicating proteinases in cancer. However, clinical studies using small molecule proteinase inhibitors have uncovered serious toxicity associated with proteinase inhibitors. Our rationale for developing protein-based inhibitors is to utilize their specificity, both as potential inhibitors themselves and as probes of potential side-effects of therapy directed at the specified targets.

We have developed a new approach for generating protein-based inhibitors that employs protein engineering to retarget the inhibitory activity of a naturally occurring inhibitor, eglin c, to the targets of interest. Molecular genetics will be used to extend the reach of traditional protein engineering by making large libraries of structural variants and then using genetic screening and selection strategies to find the best performers. Traditional protein biophysics will then be used to explore the various classes of variants and to make models for what is leading to inhibition. This information will then be used in subsequent cycles of design, construction and screening.

14. SUBJECT TERMS
Stromelysin, Eglin C, Genetic Engineering, Protein Engineering Inhibitors, Metastasis

15. NUMBER OF PAGES
8

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the US Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

MHE In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).
(no animal use this period)

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

MHE In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

MH Edell 7/21/95

PI - Signature / Date

Table of Contents

INTRODUCTION

Our objectives are to learn how to efficiently build proteins which will act as inhibitors to proteins involved in cancer metastasis. It is our expectation that such proteins can be constructed so as to be very specific for the desired target. While such inhibitors might be useful themselves as therapeutic molecules, they will certainly be useful as probes to define the issues associated with inactivating the target proteins; both the primary effects and side-effects.

The approach to be used in this project is to combine molecular genetics and protein biophysics to redirect to the target of interest the activity of a pre-existing protein which will serve as a framework onto which to mount the desired modifications. Molecular genetics will be used to extend the reach of traditional protein engineering. The idea is to make large libraries of structural variants and then use genetic screening and selection strategies to find the best performers. Traditional protein biophysics will then be used to explore the various classes of variants and to make models for what is leading to inhibition. This information will then be used in subsequent cycles of design, construction and screening.

The development cycle that we will employ to modify the wild-type eglin c into a new inhibitor is:

1. make our best design guess as to what changes will increase binding to the new target
2. construct a 'halo' of variants (~10^7) around the design
3. screen the variant library using phage display to find the best binders to the new target
4. characterize the binding classes using biophysical techniques (NMR, CD, ANS binding, etc.)
5. use the biophysical information and modeling to build hypotheses concerning binding
6. if affinities not high enough goto step 1

The simplest inhibitors bind to their targets close to or at the active site and interfere with activity simply by getting in the way and not 'letting go' of the target. These inhibitors can be designed to poke a projection into a groove in the target or enfold a pocket over a projection on the target. As our initial protein for protein engineering we wanted a molecule that was small, well mannered and for which we had some reason to think might be structurally compatible with our first set of targets. Our choice was a small proteinase inhibitor, eglin c. This protein is exceptionally stable, has no disulfide bonds, is well characterized and binds very tightly to proteins similar to our first target which is stromelysin, a proteinase implicated in metastasis. Eglin c inhibits its normal targets, serine proteinases, by binding so tightly in the Michaelis complex, that the protein cannot be raised into the transition state. A ten amino acid loop in eglin c binds within the active site groove of the native serine proteinase targets.

High affinity binding requires a sequence (binding epitope) which is compatible with the target and a set of structural constraints on that sequence which prevent it from spending much time in non-productive conformations. The engineering task, which we have set ourselves, is to replace the wild-type binding epitope with one suitable for the target and then to construct a set of new constraints to move the binding epitope into the high affinity domain. Our initial target is stromelysin and hence an appropriate binding epitope is already known, that is, a substrate sequence preferred by the proteinase. Building an inhibitor then reduces to finding a suitable series of structural constraints that can be imposed by the eglin framework on the new binding epitope.

We are pursuing two sub-lines of investigation. One is to expand our information about eglin c as a suitable framework for protein engineering and the other is to start the protein engineering with what we already know.

BODY

During this first year we have been putting together the components necessary for the engineering activities. This means taking on new personnel, constructing the various vectors we will need to carry out the protein engineering and collecting information about eglin c as a framework protein. We have:

1. Hired and trained several new personnel.
2. Constructed a new eglin c expression vector.
3. Discovered that wild-type eglin c does not function in the phage display system (our chosen screening system).
4. Embarked on a construction project to generate a circularly permuted eglin suitable for phage display.
5. Shown that a truncated form of eglin is active in the phage display system.
6. Made truncated eglin variants containing the binding epitopes for papain, collagenase, and stromelysin.
7. Made constructs for a stromelysin expressing vector.
8. Explored a new technology for studying eglin c and other proteins using a method for quantitative library assessment.

Eglin c Expression System

Prior to this project we had constructed a recombinant DNA gene for eglin c and expressed it in an expression vector called pDR720. From this strain we were able to produce about 5-30 mgs of purified eglin c per liter of culture. We have now transferred this recombinant gene via PCR to a better expression vector called pET 17b. From this vector we get 100 to 300 mgs of purified eglin c per liter. We are currently sequencing the transferred gene to make sure we have not introduced any new mutations.

The Screening System (Phage Display)

Our basic approach is to extend the reach of traditional protein engineering by constructing large libraries of structural variants around a central design concept and then using the phage display system to screen the population for binders. Hence it came as a considerable shock when we discovered that wild-type eglin c, when fused to the M13 gene III protein in the M13 particle, does not bind to a target to which the free inhibitor normally binds (e.g. subtilisin). We presume that the problem is that due to the eglin c structure attaching the phage to eglin c via it's C-terminus blocks access to the eglin c binding epitope (see Figure 1). That is, the eglin c C-terminus appears to be too close to the loop containing the residues that bind with the target proteinase.

To try to move the gene III fusion point further away from the binding epitope we made eglin c variants with various C-terminal truncations and made the M13 gene III fusions via linkers of 5 and 9 prolines. None of these eglin c variants bound to subtilisin.

An approach which is currently being pursued is to make a circularly permuted version of eglin c in which the C-terminus has been moved to the side of the protein opposite from the residues that bind with the target. We have designed an eglin variant in which the wild-type N and C termini have been joined together and new termini have been created by opening up a tight turn on the opposite side of the protein (Figure 1). The DNA sequence for this molecule has been constructed and we are in the process of inserting it into an expression vector where we can test the non-fusion form of the molecule for activity.

An alternative approach is to use a truncated form of eglin c which has already been shown to bind to its normal target and has its C-terminus located distant from the binding residues. Leatherbarrow (1991) showed that an eighteen amino acid circular peptide consisting of the amino acids in the eglin c loop containing the binding epitope and the amino acids from the underlying strands of beta sheet closed together via a cysteine bond was fully as active as native eglin c. We have determined that phage displaying this sequence bind very tightly to subtilisin. We intend to go ahead and use this truncated

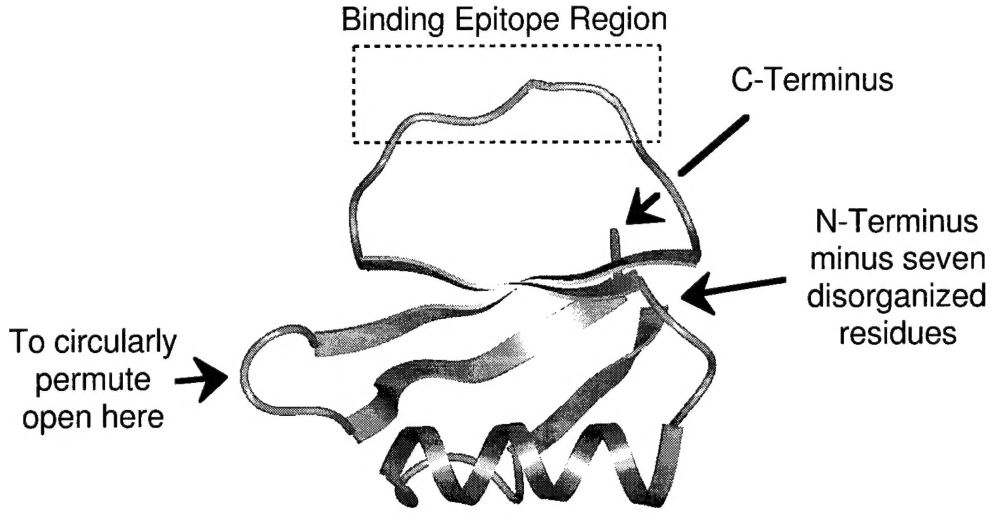

FIGURE 1. Ribbon Diagram of Eglin C. Note that attaching a phage particle to the C-terminus of eglin c might block access to the binding epitope. Our construction of a circularly permuted eglin removes the seven disorganized residues from the N-terminus, adds a four residue tight turn to connect the N- and C-terminal ends, and opens the protein to create new N- and C-termini at the point indicated in the figure.

version of eglin c for initial studies of binding but would prefer a full length molecule. High affinity binding requires that the binding epitope be held in a conformation approximating the productive conformation. While the truncated form of eglin c is apparently able to provide the necessary constraints for binding to subtilisin it is not clear whether the truncated form will be adequate to mount the constraints necessary for new binding epitopes. As a consequence we will continue with our efforts to develop a circularly permuted form for eglin so that we would have access to the whole protein for engineering .

Novel Binding Targets

Our objective is to learn how to construct protein-based inhibitors to proteins involved in the metastasis of cancer. As a model system we have chosen the proteinase stromelysin since it has been strongly implicated in metastasis, is a proteinase and hence must be able to accommodate a peptide chain within its active site. However, there is no commercial source for stromelysin so we are currently involved in constructing an expression vector containing the stromelysin sequence. A complicating feature of the construction is that while the mature sequence for stromelysin is known it is less clear what additional sequences (prepro?) are necessary for maturation of the enzyme in *E. coli.* We have constructed five variants and each produces a stable protein in *E. coli.* We now need to test these proteins for activity. For this we need an assay. The literature colorimetric assays utilize non-commercially available substrates. We think that stromelysin should work with a collagenase substrate with sufficient sensitivity for our purposes but need to test this. To do that we need some *bon fide* stromelysin and have arranged to get some from Syntex Research, Inc.. If the collagenase substrate does not work we will have senthesized for us a suitable dansylated peptide as substrate. We will then need to purify stromelysin for our affinity screening protocols. This will require preparations on the order of a few tens of milligrams and hence if our experience with pET 17b holds true should only require a few hundred milliliters. However, we will have to work out a purification protocol. Three of our constructs have a histidine tag to serve as a affinity tag for purification.

Since getting all of the stromelysin materials together will require some time we intend to work on some less direct model systems in the interim. We have chosen papain and collagenase as possible models. The idea is to initially work on two in case one gives rise to problems. Both papain and collagenase are well characterized proteins and are available from commercial sources. We have designed binding epitopes for

these two enzymes based on their cleavage specificities and constructed oligonucleotides encoding truncated eglins in which the serine proteinase binding epitopes has been replaced with ones for papain or collagenase. These sequences have now been incorporated into a plasmid vector (pGEM). The truncated eglin variant sequence bearing the collagenase binding epitope has been transferred from the plasmid vector into the m666 M13 display vector and we are now ready to test phage made with that fusion product for binding to collagenase.

Characterization of our Protein Engineering Framework
We have carried out circular dichroism spectropolarimetry (CD) for eglin c. We have determined conditions where the thermal and quanidinium chloride induced denaturation of eglin c is two-state and reversible as monitored by CD. We have determined preliminary values for the free energy, enthalphy, entropy and heat capacity of denaturation for eglin c.

We are also exploring a new approach for evaluating the rules and patterns that define the structure of a protein. It is known that two proteins can have the same fold yet share less than 15% sequence similarity. If we truly understood the rules and patterns leading to protein structure we could show that the proteins sharing little protein similarity did indeed have the same patterns. The better we understand how structures are specified in proteins in general and in eglin c in particular the more likely we will be to able to make correct predictions concerning the protein engineering changes we introduce into the protein. Preliminary results suggest the methodology works and that we can generate data in a relatively short time frame. More full-fledged experiments are underway to evaluate the rules/patterns defining the alpha helix in eglin c.

CONCLUSIONS

The project as originally defined, that is, to work out a methodology to build protein-based inhibitors against proteins involved in metastasis, remains on track. Since not all of the materials necessary for studying stromelysin are currently in hand we are currently working on the technology using papain and collagenase as model systems while we work in parallel on producing the various materials necessary to work with stromelysin.

We have learned that wild-type eglin c (the protein which we are trying to convert into a stromelysin inhibitor) does not function in the phage display screening system which we intended to use for library screening. We have created one work-around using a truncated form of eglin and are working on a more satisfying work-around involving a circularly permuted form of eglin.

We have constructed an expression vector for eglin c that increases production levels by 20 to 100 fold.

REFERENCES

Leatherbarrow, R. J. and H. J. Salacinski. Design of a small peptide-based inhibitor by modeling the active-site region of barley chymotrypsin inhibitor 2. *Biochemistry* **30** 10717-10721. (1991)